Computers in Human Behavior 135 (2022) 107370

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Sure I can code (but do I want to?). Why boys' and girls' programming beliefs differ and the effects of mandatory programming education

Una Tellhed *, Fredrik Björklund, Kalle Kallio Strand

Lund University, Sweden

ARTICLE INFO

Keywords:
Programming
Gender differences
Self-efficacy
Interest
IT
STEM

ABSTRACT

In our increasingly digitalized society, it is important to have a grasp of basic programming and therefore, programming was recently made mandatory in Swedish schools. We investigated how the mandatory programming education affects self-efficacy and interest in programming in a sample of 12-year-old children (N = 373). Data was collected at the start of the school year, directly after the education (mainly consisting of block programming), and the end of the school year. Boys had higher self-efficacy and interest than girls at all these three timepoints. Interestingly, as opposed to girls with Swedish background, girls with foreign background did not differ from the boys in programming self-efficacy. The gender difference in programming self-efficacy was completely mediated by gender differences in the children's previous mastery experience, social persuasion, and vicarious experience of computer use. Self-efficacy was strongly related to interest, but despite increasing the children's self-efficacy, programming education surprisingly reduced their interest in programming. The results regarding predictors of self-efficacy lend support to Bandura's social cognitive theory. The intersectional (gender x background) differences could be interpreted in terms of the gender-equality paradox. Possible ways of encouraging children's programming efforts and boosting the effects of interventions are discussed.

1. Introduction

Although more women have entered the STEM-sector (Science, Technology, Engineering, and Mathematics) in recent years, very few women work in IT (Information Technology) in the Western world (Nordic Council of Ministers, 2021; Su & Rounds, 2015; Swedish Agency for Economic and Regional Growth, 2020; US Bureau of Labor Statistics, 2021). Particularly jobs involving computer programming are strongly dominated by men and tech-focused educational programs, like computer science, constitute the most male-dominated in Western higher education (Department of Higher Education, 2021; National Science Foundation, 2021; Swedish Higher Education Authority, 2020; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2017).

The fact that few women are educated in computer science can be seen as problematic. In an increasingly digitalized world, the demand for people educated in computer science is rising (Organization for Economic Co-operation and Development [OECD], 2019; Swedish IT and Telecom Industries, 2021). The IT sector wants to recruit the best, regardless of gender, and many individual women could benefit from working in this sector, which has some of the best paying jobs in STEM (Nordic Council of Ministers, 2021:518; Shoffner & Dockery, 2015; Swedish IT and Telecom Industries, 2021). Also, while it may appear that women freely make the choice not to work in tech, research shows that societal gender stereotypes may cause barriers for women's self-concept development in relation to STEM which reduces their interest in STEM (e.g., Cheryan et al., 2017; Master & Meltzoff, 2020; Oldenziel, 1999; Swedish Agency for Economic and Regional Growth, 2020:3; Tellhed et al., 2021). To increase gender equity and reduce the gender segregation in the labor market, we need to find effective interventions to remove these barriers (European Institute for Gender Equality, 2017).

To strengthen future digital competence in the population, the government in Sweden recently made education in computer programming obligatory in all Swedish schools, from preschool to high school (Ministry of Education, 2017). In the current study, we follow a group of Swedish children during their school year in Grade 6, to investigate their perceptions of programming before and after they take part in the new programming teaching. We are particularly interested in testing for gender differences in programming self-efficacy and in potential sources

* Corresponding author. Department of Psychology, Lund University, Box 213, SE-221 00 Lund, Sweden.
E-mail address: una.tellhed@psy.lu.se (U. Tellhed).

https://doi.org/10.1016/j.chb.2022.107370
Received 5 April 2022; Received in revised form 17 June 2022; Accepted 20 June 2022
Available online 23 June 2022
0747-5632/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

thereof, to better understand how gender differences develop and may be altered. We will next elaborate on these concepts.

1.1. Self-efficacy and interest

The concept self-efficacy was coined by Bandura (1977; 1997) and is defined as the belief in one's capabilities to exercise certain tasks, such as learning computer programming. It is closely related to "ability beliefs" which refers to a more general belief in one's ability, such as having technical skills (Wigfield & Eccles, 2000). For reasons of simplicity and to cohere with our measurements' terminology, we mainly use the term self-efficacy in this article.

In his social cognitive theory, Bandura (1986) stated that having self-efficacy is important for making choices, taking on challenges, and persevering when facing difficulties. Research has shown that self-efficacy better predicts educational interest than ability tests do (Bandura, 1997; Brown et al., 2008; Lent & Brown, 2019). It has also been shown that girls may underestimate their ability in STEM-subjects, such as math or technology, while boys sometimes overestimate it (Bench et al., 2015; Hackett, 1995; Watt, 2010). Much research has shown that gender differences in STEM-interest and STEM-choice is closely related to gender differences in self-efficacy and ability beliefs for STEM-subjects (e.g., Eccles & Wang, 2016; Hackett, 1995; Lent & Brown, 2019; Master & Meltzoff, 2020; Tellhed et al., 2017, 2018; Rottinghaus et al., 2003; Watt, 2010). This finding supports well-established theories such as social cognitive career theory (SCCT, Lent et al., 1994), Eccles's expectancy value theory (EEVT; Eccles, 1987, 1994) and the recent stereotypes, motivation, and outcomes model (STEMO; Master & Meltzoff, 2020). SCCT expands on Bandura's social cognitive theory and states that career interests are influenced by self-efficacy and outcome expectations (beliefs about what we can get) in relation to career opportunities (Lent et al., 1994; Lent & Brown, 2006). Similarly, EEVT states that career choice is influenced by the individual's expectations of success (which in turn relates to ability beliefs) and by the perceived task value of different career options (Eccles, 1987, 1994). STEMO (Master & Meltzoff, 2020) builds upon EEVT and expands on it, predominately by emphasizing the role of sense of belonging (fitting in) for career interest and choice. All three theories state that gender bias may affect career interest and choice, but particularly EEVT and STEMO place a specific weight on the role of gender stereotypes for factors related to gender differences in career interest and choice. Recently, two studies in Sweden, where the current data is collected, with large representative participant samples, showed that the strongest mediator of gender differences in STEM-interest and STEM-choice was self-efficacy/ability beliefs, and that social belongingness expectations also contributed to explain the difference (Tellhed et al., 2017; 2018). These results thus supported both SCCT, EEVT and STEMO.

1.2. Programming self-efficacy and the gender equality paradox

Although many have studied gender differences in STEM self-efficacy, there is very little research on gender differences in *programming* self-efficacy and how it relates to children's programming interest. One previous study in Hong Kong found no gender difference in 14-year-olds' programming self-efficacy although boys had greater programming interest than girls (Kong et al., 2018). Another study with 6-year--old children in the US found that boys rated programming as more fun than girls and rated higher than girls on the question "how good are you with robots" (Master et al., 2017). There are also only a few studies on gender programming self-efficacy in adult students taking programming classes in Turkey, Taiwan, and the US, and all but one found no gender difference (Korkmaz & Altun, 2014; Lin, 2016; Ramalingam & Wiedenbeck, 1998, but see Askar & Davenport, 2009, that found a gender difference). Further, Tsai (2019) showed a promising large increase in programming understanding from a programming intervention aimed at university students in Taiwan, and the effect was larger for students with low or moderate self-efficacy. However, this study did not report gender differences.

In the current study, we investigate programming self-efficacy in a sample of 12-year-old children in Sweden and in contrast to the previous studies, we expect to find a gender difference. We base this prediction on the recently discovered "gender equality paradox", which has shown that countries that score highly on gender equality indexes, like Sweden does (European Institute for Gender Equality, 2021), tend to show larger self-rated psychological gender differences than other countries (e.g., Stoet & Geary, 2018). Note that this paradox regards self-rated constructs, like career interests, while ability tests generally show gender similarity (Else-Quest et al., 2010; Hyde et al., 2019, 2005; Zell et al., 2015). In the last decade, many immigrants have come to Sweden from countries which score lower on gender-equality indexes than Sweden does, and where many more women study STEM (Huyer, 2015). Therefore, we also aim to explore possible intersectional effects of participant ethnicity with participant gender on programming self-efficacy and programming interest in this study, if the demographic factors in the participant sample will allow for it.

1.2.1. The role of gender stereotypes

A study recently showed that the gender equality paradox is explained by the circumstance that people in more gender-equal countries tend to have stronger gender-stereotypes than others (Breda et al., 2020). Interestingly, although male primacy is generally rejected in more gender-equal countries, gender-essentialist beliefs are still strong, where many believe that men and women are born psychologically very different, and therefore for example fit different types of occupations, (Knight & Brinton, 2017). There is empirical support that the specific stereotype that women are less technically skilled than men is very strong in Sweden (Master et al., 2017; Mellström, 2004, 2009; Tellhed et al., 2021). Drawing on EEVT (Eccles, 1987, 1994), STEMO (Master & Meltzoff, 2020) and stereotype threat theory (STT; Steele & Aronson, 1995; Steele, 1997) this stereotype could hinder girls from developing robust self-efficacy for technology, including programming self-efficacy. To exemplify, when stereotypes portray a group as lower in ability than others, it can trigger stereotype threat in negatively stereotyped group members (Spencer et al., 2016; Steele, 1997; Steele & Aronson, 1995). Stereotype threat has been shown to confuse the individual's performance judgments, which could make them more susceptible to internalize negative ingroup stereotypes in their self-concept and hinder their self-efficacy development (Aronson & Inzlicht, 2004; Tellhed & Adolfsson, 2018).

1.2.2. Physiological arousal

Stereotype threat triggers *physiological arousal*, which Bandura (1977, 1997) has identified as one source of self-efficacy development in an individual. According to his social cognitive theory, associating stress or anxiety with a domain interferes with developing strong self-efficacy in it. Relatedly, previous research has shown that anxious feelings when working with computers is more common in women than men (e.g., Cooper, 2006). Because most Swedish children use computers daily, they are likely not anxious when using them. However, it is possible that Swedish girls feel more nervous than boys when asked to perform more technically advanced computer tasks, where they could be reminded of the stereotype that girls are perceived as less technically skilled than boys. Possibly, this triggered arousal could be one source of Swedish girls hypothesized weaker programming self-efficacy than boys.

1.3. Sources of gender differences in self-efficacy

Apart from physiological arousal, Bandura (1977, 1997) theorized that mastery experience, vicarious experience, and social (a.k.a. verbal) persuasion are important sources for self-efficacy development. Mastery experience is the perception of succeeding or failing during performance

experience in a domain, vicarious experience is the perception of identifiable role models as successful, and social persuasion is the encouragement, or discouragement, one receives from others when engaging in a task. Previous research has shown gender differences in sources of self-efficacy for some gender-stereotypical domains like math or reading (Usher & Pajares, 2008). In a study of adult university students in Taiwan, Lin (2016) did not find gender differences in sources for programming self-efficacy and relatedly not in programming self-efficacy. In the current study, we expand previous research to investigate sources of gender differences in Swedish children's programming self-efficacy and expect to find gender differences in all four proposed sources.

1.3.1. Mastery experience

Bandura believed that mastery experience is the most powerful source of self-efficacy and meta-analytical findings support this (Usher & Pajares, 2008), although the importance of different sources varies over time and sometimes gender (Phan & Ngu, 2016). Further, tapping into several sources when designing a self-efficacy intervention increases its impact (Unrau et al., 2018).

Children can have mastery experience both during and after school. Interestingly, men in STEM often claim to have developed their STEM-interest in their free time, while women in STEM claim to have developed it in school (Maltese & Tai, 2010). Note that programming has only very recently been taught systematically in schools in Sweden, and we also suspect that few children program computers in their free time. However, mastery experience with computers often starts with computer games, which boys play more frequently than girls (Cooper, 2006:; Master et al., 2017; Swedish Media Council, 2021; Vermeulen et al., 2017). If boys have more mastery experience with computers through gaming, it could perhaps help them develop a stronger sense of programming self-efficacy than girls.

1.3.2. Vicarious experience

At least in Western countries, the prototypical computer programmer is perceived as a white man with nerdy interests, weak social skills, and a brilliant mind (Cheryan et al., 2011; Starr, 2018). This stereotype contrasts with norms for femininity and can deter some girls' interest and sense of belonging in IT (Cheryan et al., 2011; Starr, 2018). The lack of plentiful, identifiable, same gender role models thus could be another obstacle for girls' programming self-efficacy development (Dennehy & Dasgupta, 2017; Olsson & Martiny, 2018). Congruently, we expect that Swedish boys have had more instances of vicarious experience in programming than girls, which could have benefited boys' self-efficacy development.

1.3.3. Social persuasion

Lastly, previous research has shown that girls are less encouraged (and occasionally even discouraged) from pursuing STEM-careers, as compared to boys (Betz & Hackett, 1981; Cadaret et al., 2017; Master & Meltzoff, 2020). This could relate to adults, including teachers, implicitly associating STEM with men and unintentionally conveying their bias to children, which may hinder girls' programming self-efficacy development (Cadaret et al., 2017; Ertl, Luttenberger, & Paechter, 2017; Master & Meltzoff, 2020; Miller et al., 2015; Tellhed et al., 2021). Receiving encouragement appears especially important for negatively stereotyped groups, which may relate to the above-mentioned importance of school for girls' STEM interest development (see Shoffner & Dockery, 2015, for a review). One study showed that women in STEM perceived that encouragement from role models like STEM-teachers was their main source of STEM self-efficacy, while men in STEM perceived mastery experience as their main source (Zeldin et al., 2008).

1.4. The effect of programming education

In this study, we expand previous research by investigating if children's programming self-efficacy and interest changes after taking part in the programming education, which is now mandatory in all Swedish schools. We measure children's perceptions when they begin Grade 6, then again after they have had the programming education, and follow up at the end of the school year.

Being taught programming in school represent a form of mastery experience, which should raise children's programming self-efficacy, if they felt sufficiently successful during the exercise (Bandura, 1977, 1997). We have not seen a similar evaluation of implementation of programming education in elementary school in the literature, but one study found that when 6-year-old US children programmed robots in a laboratory, it strongly raised girls' self-efficacy for robots and moderately raised their programming interest (Master et al., 2017). Boys' perceptions were not raised by this intervention, which could be due to a ceiling effect (Master et al., 2017). There are also a few studies which has qualitatively evaluated kindergarten student's enjoyment of similar programming tasks (Fessakis et al., 2013; Kalelioğlu & Gülbahar, 2014; Keren & Fridin, 2014; Rogozhkina & Kushnirenko, 2011).

Programming is not a graded subject in Swedish elementary schools and there is no official examination form, but we will ask teachers to judge each child's programming performance in class, as compared to their classmates. Note that this is not an objective test of the children's programming ability, but previous studies tend to show no gender difference among students in computer science (e.g., Lin, 2016). Since encouragement have been shown to matter especially for girls' self-efficacy development (see Shoffner & Dockery, 2015, for a review), we expect that children, and perhaps especially girls, who the teacher perceives as performing better, gets a stronger self-efficacy boost than others. Lastly, drawing on SCCT, EEVT, and STEMO, increases in programming self-efficacy should relate to increases in programming interest.

1.5. Hypotheses

To summarize, the hypotheses are:

1. a) Girls have lower programming self-efficacy than boys and b) lower programming interest than boys.
2. A gender difference in programming self-efficacy is mediated by previous a) mastery experience, b) vicarious experience, c) social persuasion, and d) physiological arousal.
3. A gender difference in programming interest is mediated by self-efficacy.
4. Taking part in programming education increases children's programming self-efficacy.
5. Increases in programming self-efficacy over time is related to a) self-rated performance in programming class, b) teacher-rated performance in programming class.
6. Increases in self-efficacy after programming education relates to increases in programming interest.
7. We explore intersectional effects of participant ethnicity with participant gender on a) programming self-efficacy and b) programming interest.

2. Material and methods

2.1. Project Description and Compliance with Ethical Standards

The data presented here is part of a larger project funded by the Swedish Research Council for Health, Working Life, and Welfare (grant #2019-00334) and approved by the Swedish Ethical Review Authority (#2020-01606).

We informed the school pupils that participation in the study was voluntary and collected informed consent from both the children and their parents. We provided each participant with a personal code that they entered when taking the three questionnaires, so that their

responses could be connected. After finishing the data collection, we destroyed the code key, which connected codes to person names. We gave each pupil an ice cream gift card when they filled out the last survey.

2.2. Study design and procedure

The study design was longitudinal with repeated measures at three timepoints. Timepoint 1 (T1) was a baseline measure at beginning of the school year (fall semester 2020). Timepoint 2 (T2) was soon after the pupils had their programming activities in school. Timepoint 3 (T3) was a post-measure taken at the end of the school year (the end of the spring semester 2021), to test if effects from T2 had persisted or diminished over time. The timing of the programming education, and thereby the T2 measure, varied somewhat between school classes. A control group could not be included in the study design, since programming education is now mandatory for all students in Sweden (Ministry of Education, 2017).

We carried out all testing of the pupils during school hours and the teachers completed their performance ratings of the pupils at their own convenience.

2.2.1. The programming education

Programming education has been mandatory in Swedish schools since 2018. The curriculum for Grade 4–6 states that the core content in the teaching should concern "How algorithms can be created and used in programming", "Programming in visual programming environments", and "Controlling pupils' own constructions or other objects by means of programming" (Swedish National Agency for Education, 2018). There are no directions regarding which specific exercises that it should involve and there was recently a call for more appropriate educational material (Official reports of the Swedish Government, 2021:70). Teachers tend to find general phrases, like in the curriculum, too vague for determining what the pupils need to learn (Vinnervik, 2022a). It should be noted that general phrasings are not uncommon in Swedish curriculums, to provide opportunity for teachers with adequate expertise to design exercises the way they find best, and it is not always considered a problem. But since no less than 70% of teachers appointed to teach programming (mostly math teachers) state that they lack confidence in teaching it (The National Union of Teachers in TheNational Union of Teachers in Sweden, 2020) the demand for clearer directions is understandable. In addition to this, the curriculum encourages collaboration between different teachers (not only those teaching technology and math). Such ambitious goals in combination with the perception of having too little time and resources for preparation has brought some complaints from teachers, as well as calls for clearer guidelines (Official reports of the Swedish Government, 2021:70, Vinnervik, 2022b).

This means that the programming education we investigate the effects of is still developing and varies between schools and school classes. Those teaching the pupils in the current study reported focusing mainly on block programming, as well as reasoning about what programming is and the purpose of it. They typically devoted 5–7 lessons (each 40–60 min long) to programming education. The majority used scratch (http://scratch.mit.edu) for the exercises, but some used Lego Mindstorms or micro:bit (https://microbit.org). A slight majority of the teachers had taken a preparatory programming education course, either at university or their workplace.

2.3. Participants

A total of 407 children in Grade 6 from 15 different schools in the south of Sweden were recruited. Thirty-four pupils did not participate at T1 and were excluded; thus, 373 pupils (Mean age at T1 = 11.87, *SD* = 0.44) were included in the analysis. Of these pupils, 313 participated at T2 and 323 participated at T3. The vast majority (96.8%) self-identified with their legal gender (which in Sweden is woman or man) which we used for gender difference analyses. The gender balance was even, with 53.9% boys and 46.1% girls. Official statistics in Sweden registers ethnicity as "Swedish" contra "Foreign background" where the latter is defined as being born outside of Sweden or having both biological parents born outside of Sweden (Statistics Sweden, 2002). With this definition, 27% of the pupils had a foreign background. As a reference point, approximately 20% of the population in Sweden has a foreign background (Migrationsinfo, 2020).

2.3.1. Programming and gaming experience

We asked the participants to estimate on a five-point scale (Never/Seldom/Once a week/A few times a week/Every day) how much time they spend on programming in their leisure time, and how much time they spend on computer games. Table 1 shows that the children had little previous experience programming computers, but the boys had slightly more experience than the girls. Regarding computer gaming, most boys rated that they play daily, while the girls gamed much more seldom than the boys (see Table 1). These factors were only collected for descriptive purposes in this study, but Table A1 in the Supplemental Material shows their correlation with other measures.

2.4. Measures

We collected the data through Qualtrics. We subjected the test items to back translation, such that all materials were in Swedish. The measures presented in this article represent part of the larger project (see Project Description and Compliance with Ethical Standards) and we will present other results elsewhere. The dataset and code are found in the open science data repository, (https://osf.io/4h5pu).

2.4.1. Programming self-efficacy

We measured programming self-efficacy with a 7-item scale (e.g., "I am good at programming" and "I have confidence in my ability to program") that we adapted from the Kong et al. (2018) scale but revised to also include two reversed items (inspired by Tek et al., 2018). It had a 5-point response scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The scale has previously been confirmed with regards to convergent and content validity (Kong et al., 2018). Cronbach's alpha was .81 at T1, 0.85 at T2, and 0.86 at T3.

2.4.2. Sources of self-efficacy

There were eight items in the sources of self-efficacy measure, which we adapted from Usher and Pajares (2008; 2009). Since we expected that the children would have very little previous programming experience (which was corroborated, see section 2.2.1. above), we decided to measure previous experience with the more abstract concept "computers". There were two items each for the four sources: mastery experience (e.g., "I have always been good at computers"), vicarious experience (e.g., "When meeting kids who know computers better than me, I think that I can get better too"), social persuasion (e.g., "People have told me that I have talent for computers", and psychological arousal (e.g., "I start to feel stressed out when beginning to do something technical with computers"). Responses were provided on the same 5-point scale as for self-efficacy beliefs. A previous validation study found the original scale to be psychometrically sound with strong content validity and criterion validity (Usher & Pajares, 2009). Cronbach's alpha at T1 was 0.74 for mastery experience, 0.52 for vicarious experience, 0.87 for social persuasion and .61 for physiological arousal.

2.4.3. Programming interest

We measured programming interest with the two items "How interesting do you think programming is?" and "How interesting do you think it will be to have programming in school?", rated on a 5-point scale ranging from 1 (*not at all interesting*) to 5 (*extremely interesting*). Cronbach's alpha was .83 at T1, 0.85 at T2, and 0.85 at T3.

Table 1
Descriptive Statistics and Gender Comparisons.

Variable	Total		Girls		Boys		Gender comparisons			
	n	*M* (*SD*)	*n*	*M* (*SD*)	*n*	*M* (*SD*)	*df*	*t*	*d*	95% CI
Self-Efficacy										
Timepoint 1	373	3.24 (0.74)	172	3.08 (0.70)	201	3.37 (0.74)	366.92	−3.85	−0.40	[−0.44, −0.14]
Timepoint 2	313	3.44 (0.77)	151	3.27 (0.76)	162	3.60 (0.74)	307.77	−3.90	−0.44	[−0.50, −0.16]
Timepoint 3	323	3.37 (0.84)	149	3.19 (0.75)	174	3.52 (0.88)	321.00	−3.65	−0.41	[−0.51, −0.15]
Programming interest										
Timepoint 1	373	3.46 (0.96)	172	3.30 (0.92)	201	3.60 (0.97)	367.43	−3.04	−0.31	[−0.49, −0.11]
Timepoint 2	313	3.34 (1.03)	151	3.17 (1.02)	162	3.50 (1.01)	309.01	−2.89	−0.33	[−0.56, −0.11]
Timepoint 3	323	3.17 (1.10)	149	2.98 (1.03)	174	3.34 (1.14)	320.18	−2.93	−0.33	[−0.59, −0.12]
Mastery experience	373	3.32 (1.01)	172	2.98 (0.97)	201	3.61 (0.95)	359.06	−6.31	−0.66	[−0.83, −0.43]
Vicarious experience	373	3.57 (0.94)	172	3.38 (0.92)	201	3.72 (0.92)	362.31	−3.52	−0.37	[−0.53, −0.15]
Verbal persuasion	373	2.54 (1.26)	172	2.20 (1.17)	201	2.84 (1.25)	368.14	−5.08	−0.53	[−0.89, −0.39]
Physiological arousal	373	1.76 (0.89)	172	1.83 (0.93)	201	1.70 (0.84)	347.31	1.31	0.14	[−0.06, 0.30]
Free-time computer gaming	373	3.54 (1.37)	172	2.70 (1.29)	201	4.26 (0.96)	310.47	−13.06	−1.37	[−1.80, −1.33]
Free-time programming	373	2.01 (0.93)	172	1.78 (0.83)	201	2.21 (0.97)	370.99	−4.66	−0.48	[−0.62, −0.25]
Self-rated performance	268	3.34 (0.91)	128	3.10 (0.80)	140	3.56 (0.95)	264.52	−4.33	−0.53	[−0.67, −0.25]
Teacher-rated performance	318	3.43 (1.01)	147	3.33 (0.92)	171	3.51 (1.08)	315.99	−1.56	−0.17	[−0.40, 0.05]

Note. t = Welch's t-test.

2.4.4. Self-rated and teacher-rated performance

After the programming education, the pupils rated themselves on the item "How well do you think you performed on the programming exercises, as compared to your classmates?" on a 5-point scale ranging from 1 (*Among the worst in class*), over 3 (*Like most in class*) to 5 (*Among the best in class*). The teacher responsible for the programing education rated each pupil's performance in class on the same above scale.

2.5. Data analysis

The data was analyzed in R version 4.1.2 (R Core Team, 2021). All categorical variables used in the analyses were coded using Helmert contrasts (e.g., Gender coded as −1 for boys, 1 for girls). Hypothesis testing for the longitudinal data was carried out using linear mixed effects modeling. We fit linear mixed effects models using the lme4 package (Bates et al., 2015) with p-values generated using the lmerTest package (Kuznetsova et al., 2017). All models were tested to ensure no violations of the assumptions of linear mixed effects modeling, namely normality, linearity, multicollinearity, homoskedasticity, and influential outliers.

3. Results

3.1. Gender differences

Table 1 presents the descriptive statistics for the observed variables at the three timepoints, along with independent samples t-tests comparing girls with boys as well as total mean values. The results revealed gender differences across most of the measures. In support of Hypothesis 1A and 1B, compared to boys, girls reported significantly lower programming self-efficacy and programming interest across all three timepoints. Girls also reported lower mastery experience, social persuasion, and vicarious experience, but there was no significant gender difference in physiological arousal. The effect sizes ranged from small to medium. Although the boys' self-rated programming performance was higher than the girls', there was no gender difference in the teacher-rated performance. The children's self-ratings and their teacher's ratings of their programming ability was significantly correlated for both boys and girls, $r = 0.35$, $p < .001$.

3.2. Mediation analyses of gender differences

To test Hypothesis 2A-2D we investigated whether the gender differences in self-efficacy could be explained by gender differences in sources of self-efficacy. We conducted a mediation analysis using structural equation modeling (SEM) in the lavaan package (Rosseel, 2012). A direct path was added between Gender and Self-Efficacy, as well as indirect paths via each Mastery Experience, Vicarious Experience, and Social Persuasion. Physiological Arousal was not included as a mediator in the model as we did not find any gender differences in this variable (see Table 1), nor was it related to self-efficacy (see Supplemental Material). Lastly, to investigate Hypothesis 3, we also included direct paths from Self-Efficacy and Gender to Programming Interest. Fig. 1 illustrates the structural model with standardized path coefficients (see Table A2 in the Supplemental Material for all parameter estimates). The model did not pass the chi-square test but showed overall good fit based on approximate fit indexes, χ^2 (3) = 12.83, $p = .005$, CFI = 0.985, RMSEA = 0.094 (90% CI 0.045–0.149), SRMR = 0.032. In line with our predictions, there was no direct effect of gender on self-efficacy in the model; instead, the relationship was fully mediated by the three sources of self-efficacy in the model, with the strongest relationship via Mastery Experience. The indirect effects of gender via the mediators together explained 3.6% of the variance in self-efficacy. Moreover, the sources of self-efficacy explained 16.1% of the variance in self-efficacy over and beyond the indirect effect of gender. Lastly, self-efficacy fully mediated the relationship between gender and programming interest. Gender was not directly related to programming interest in the model, but the total indirect effect via self-efficacy (and thus also via the sources of self-efficacy) was significant albeit weak, explaining 1.1% of the variance in interest. Self-efficacy was strongly related to interest, explaining 28.8% of the variance over and beyond the indirect effect of gender.

3.3. The effect of the education on self-efficacy

Next, we tested whether the programming education impacted the children's programming self-efficacy (Hypothesis 4). In Model 1, self-efficacy was the dependent variable and Time was the predictor. Because the pupils were measured on three timepoints in total, we included Time as a quadratic term in addition to the linear effect, to account for potential nonlinear effects over time. Moreover, we included a random intercept and slope for each pupil as well as a random intercept for each school. The model is presented in Table 2. As predicted, there was a significant but weak linear positive change over time in programming self-efficacy. Moreover, a significant positive correlation ($r = 0.27$) between the random intercept and the random slope for each pupil indicated that pupils with higher initial self-efficacy showed a stronger self-efficacy increase over time compared to pupils with lower initial self-efficacy. The quadratic effect of time was also significant; self-efficacy increased between T1 and T2 but regressed slightly between T2 and T3. The variance explained by the model could largely be attributed

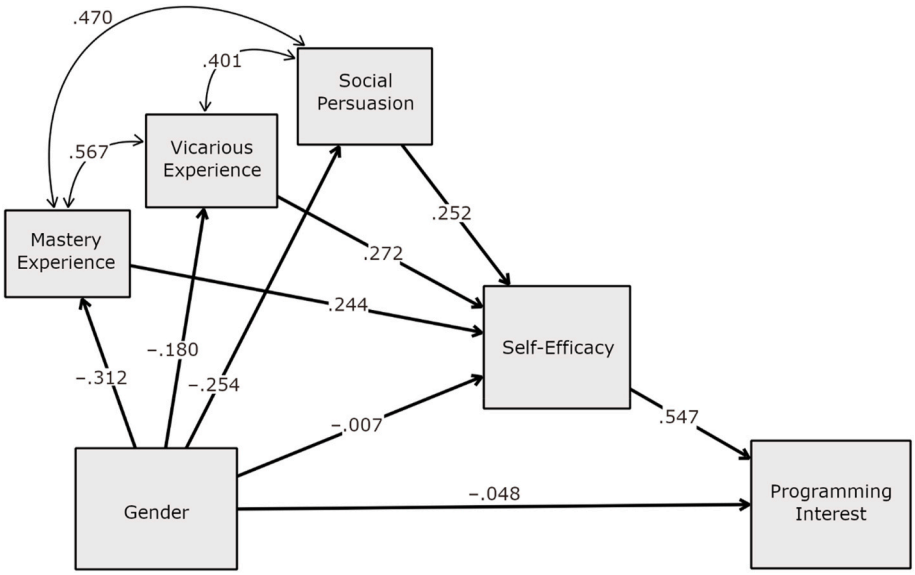

Fig. 1. Structural Model of Direct and Indirect Effects on Programming Self-Efficacy and Interest (Standardized Path Coefficients).

Table 2
Summary Statistics for Model 1.

Predictor	β	*p*	95% CI	R^2
$Time_{Linear}$	0.07	<0.001	[0.03, 0.11]	0.005
$Time_{Quadratic}$	−0.08	<0.001	[–0.11, −0.05]	0.006

Note. Model $R^2_m = 0.011$, $R^2_c = 0.778$.

to the random intercept for each pupil.

3.3.1. Relation to performance ratings

Next, we tested whether the effect of the education on self-efficacy related to the pupils' self-rated performance in class and to their teacher's ratings of their performance (Hypothesis 5A and 5B). We also explored if participant gender affected these relations. We built a new linear mixed effects model (Model 2) with self-efficacy as the independent variable and linear time, quadratic time, gender, self-rated performance, and teacher-rated performance as predictors. We included three-way interactions between gender, linear time, and self- and teacher-rated performance, respectively. The model output for the relevant effects is provided in Table 3. The results revealed the same linear and quadratic effects of time as found in Model 1. Both self- and teacher-rated performance were significant positive predictors of self-efficacy, with self-rated performance being the strongest predictor. No significant main effect of gender was observed, nor were there any two-way interactions. There was however a significant three-way interaction for Gender × Linear Time × Teacher-Rated Performance as well as for Gender × Linear Time × Self-Rated Performance. To further understand these interactions, we split the analysis by gender and investigated the two interaction effects (i.e., Linear Time × Teacher-Rated Performance and Linear Time × Self-Rated Performance) separately for boys and girls. For boys, the over-time change in self-efficacy was contingent on their self-rated performance ($p = .009$), with high ratings associated with higher self-efficacy increases and low ratings associated with decreases in self-efficacy. By contrast, boys' self-efficacy change over time was not influenced by the teacher-rated performance ($p = .788$), although teacher-rated performance was positively associated with self-efficacy at T1 ($p = .003$). For girls, the pattern was reversed. Girls' over-time changes in self-efficacy were not influenced by their self-rated performance ($p = .492$), although higher self-rated performance was associated with higher self-efficacy at T1 ($p < .001$). Instead, girls' over-time self-efficacy changes were contingent on teacher-rated performance ($p = .009$), with high teacher-rated performance associated with higher increases in self-efficacy and low teacher-rated performance associated with over-time decreases in self-efficacy.

Table 3
Summary Statistics for Model 2.

Predictor	β	*p*	95% CI	R^2
$Time_{Linear}$	0.06	0.004	[0.02, 0.11]	0.005
$Time_{Quadratic}$	−0.09	<0.001	[–0.13, −0.06]	0.011
Gender (base = boys)	−0.06	0.259	[–0.15, 0.04]	0.004
Self-rated performance (SRP)	0.47	<0.001	[0.37, 0.57]	0.202
Teacher-rated performance (TRP)	0.10	0.039	[0.01, 0.19]	0.012
Gender × $Time_{Linear}$	−0.03	0.091	[–0.08, 0.01]	0.002
Gender × SRP	−0.01	0.843	[–0.11, 0.09]	<0.001
Gender × TRP	−0.09	0.060	[–0.19, 0.00]	0.010
$Time_{Linear}$ × SRP	0.03	0.243	[–0.02, 0.07]	0.001
$Time_{Linear}$ × TRP	0.04	0.095	[–0.01, 0.08]	0.002
Gender × $Time_{Linear}$ × SRP	−0.05	0.028	[–0.10, −0.01]	0.003
Gender × $Time_{Linear}$ × TRP	0.05	0.041	[<0.01, 0.09]	0.003

Note. Model $R^2_m = 0.319$, $R^2_c = 0.782$.

3.4. The effect of the education on programming interest

Next, we investigated whether increases in self-efficacy over time would relate to increases in programming interest over time (Hypothesis 6). We conducted a mediation analysis with bootstrapping ($n = 10{,}000$) using the mediation package (Tingley et al., 2014) to investigate the extent to which the over-time changes in programming interest could be explained by over-time changes in self-efficacy, using data from T1 and T2. We included a random intercept for each pupil and cluster-centered self-efficacy in two steps: first within schools, then within pupils. The results are illustrated in Table 4. There was a significant direct effect of time on programming interest (Step 1), in that pupils' interest unexpectedly decreased over time following the programming education. There was also a positive indirect effect of time via self-efficacy, in that over-time changes in self-efficacy were positively related to programming interest; time was a positive predictor of self-efficacy (Step 2), and self-efficacy positively predicted programming interest (Step 3). This indirect effect via self-efficacy explained 65.4% of the relationship between time and interest. Taken together, although programming interest

Table 4
Mediation Analysis of the Effect of Time on Programming Interest via *Self-Efficacy*.

Model	b	*p*	95% CI	R^2
Step 1				
Time → Interest	−0.08	0.031	[−0.16, −0.01]	0.003
Step 2				
Time → Self-Efficacy	0.15	<0.001	[0.11, 0.19]	0.069
Step 3				
Time → Interest	−0.14	<0.001	[−0.22, −0.07]	0.007
Self-Efficacy → Interest	0.38	<0.001	[0.24, 0.53]	0.017

decreased over time following the programming education, an increase in programming self-efficacy over time counteracted this decrease.

3.5. Exploratory intersectional analyses

Lastly, we conducted exploratory intersectional analyses for programming self-efficacy and interest at T1 based on Participant Ethnicity and Gender, as approximately a quarter of the pupils were coded as having a foreign background (Hypothesis 7A and 7B). We conducted two 2 × 2 ANOVAs with Gender and Ethnicity as independent variables and programming self-efficacy and interest as dependent variables in the respective analyses. The descriptive statistics for the analyses are provided in Table 5. We used the Anova() function in the car package (Fox & Weisberg, 2019) with Type III sums of squares.

In both analyses, the same significant gender differences were found as reported in the previous analyses. There was also a significant effect of Ethnicity in both analyses, where pupils with a foreign background reported both higher self-efficacy, $F\,(1,\,363) = 10.42$, $p = .001$, $\eta_p^2 = 0.028$, and higher programming interest, $F\,(1,\,363) = 7.99$, $p = .005$, $\eta_p^2 = 0.022$, compared to pupils with a Swedish background. Interestingly, for self-efficacy there was a significant interaction between Gender and Ethnicity, $F\,(1,\,363) = 5.62$, $p = .018$, $\eta_p^2 = 0.015$, which is illustrated in Fig. 2. Tukey's HSD post hoc tests revealed that while there was no gender difference for pupils with a foreign background ($p = .999$) nor any effect of ethnicity among boys ($p = .920$), girls with a Swedish background had significantly lower self-efficacy than all three other groups ($ps < .002$). No such interaction was found for programming interest, $F\,(1,\,363) = 2.83$, $p = .094$, $\eta_p^2 = 0.008$.[1]

4. Discussion

In this study, we investigated gender differences in programming self-efficacy in a sample of Swedish children and related it to Bandura's four proposed sources of self-efficacy (Bandura, 1977, 1997). We also tested if the children's programming self-efficacy was affected by taking part in programming education, which was recently implemented in Swedish schools, and tested if changes in self-efficacy related to changes in programming interest.

4.1. Sources of gender differences in programming self-efficacy

As predicted, the results showed that boys rated their programming self-efficacy higher than girls. This result contrasts with a previous study with similar-aged children in Hong Kong (Kong et al., 2018), which did not find a gender difference in programming self-efficacy. We will elaborate on this potential cross-country difference in the intersectional analysis part below but will first discuss possible origins of the gender difference in Sweden.

The gender difference in programming self-efficacy was completely mediated by gender differences in three of Bandura's (1977, 1997) proposed sources of self-efficacy. The mediators were previous mastery experience, social persuasion, and vicarious experience in relation to computers, which were all moderately related to self-efficacy. Mastery experience is generally considered the most important source of self-efficacy (Bandura, 1977, 1997; Usher & Pajares, 2008) and previous mastery experience with computers showed the largest gender difference in this study. In addition to having more mastery experience with computers than girls, the boys had also experienced more encouragement and praise for their talent with computers (social persuasion) as well as had more experience with influential role models than girls (vicarious experience).

However, contrary to expectations, there was no gender difference in the children's feelings of stress or tenseness when solving technical computer issues (physiological arousal) and this factor was not related to programming self-efficacy in this study. Drawing on stereotype threat theory (Steele, 1997; Steele & Aronson, 1995), we expected girls to have experienced more negative stress in relation to technical computer issues than boys, since women are stereotyped as having lower technical ability than men in Sweden (Mellström, 2009; Tellhed et al., 2021). Fortunately, the children reported very low levels of perceived stress in relation to handling technical computer issues, which perhaps relates to most Swedish children using computers from a very early age (Swedish Media Council, 2021). Although this is good news, we encourage further investigation of physiological arousal as a potential source of gender difference in programming self-efficacy. We did not measure arousal in relation to computer programming specifically, as we expected the children to have little experience of programming, which was supported by the results. As children get more experience with computer programming through the continuous education in school, it will be possible to measure sources of self-efficacy in relation to computer programming specifically. We encourage such research as predictors are typically more valid when measured at a similar level of abstraction as the outcome variable (e.g., Meissner et al., 2019).

That three of Bandura's proposed sources of self-efficacy completely mediated the gender difference in programming self-efficacy in this study implies that to increase girls' self-efficacy we should allow for girls to have more experiences with computer programming, encourage their efforts and allow for identifiable programming role models also for girls. Increasing positive sources of programming self-efficacy for girls could be applied in school settings, which brings us to the result of how Sweden's new programming education impacted the children's self-efficacy in this study.

4.2. Programming education increased self-efficacy

Learning programming in school is a form of mastery experience and we were therefore interested to learn how the school reform, where programming is now taught in all Swedish schools, may affect children's self-efficacy. The results showed that taking part in the handful of programming lessons given during Grade 6 raised both boys' and girls' self-efficacy as compared to their ratings at the beginning of the school year. Although the raise in self-efficacy was small, and slightly reduced at the follow up measure at the end of the school year, it is promising that a few programming lessons could raise self-efficacy. The result supports Bandura's social cognitive theory (1977, 1997) and implies that when given more programming lessons, the students' self-efficacy could rise even further. We encourage more research in educational settings, to see how self-efficacy develops in relation to repeated programming education.

4.2.1. The education and gender

The teachers rated the boys' and the girls' programming skills similarly. This fits with previous research that tends to find gender similarity in most psychological factors (Else-Quest et al., 2010; Hyde et al., 2019, 2005; Zell et al., 2015) including programming

[1] The model with programming interest violated the assumptions of normality. We therefore conducted a robust two-way ANOVA using the t2way() function in the WRS2 package (Mair & Wilcox, 2020). This yielded the same pattern of results as the original 2 × 2 ANOVA, lending support to the original results.

Table 5
Descriptive Statistics for the 2 × 2 ANOVAs.

Variable	Swedish Background				Foreign Background			
	Girls		Boys		Girls		Boys	
	n	*M* (*SD*)	*n*	*M* (*SD*)	*n*	*M* (*SD*)	*n*	*M* (*SD*)
Self-Efficacy	126	2.96 (0.68)	142	3.36 (0.75)	44	3.44 (0.67)	55	3.44 (0.73)
Programming Interest	126	3.16 (0.85)	142	3.57 (0.97)	44	3.66 (1.02)	55	3.70 (0.93)

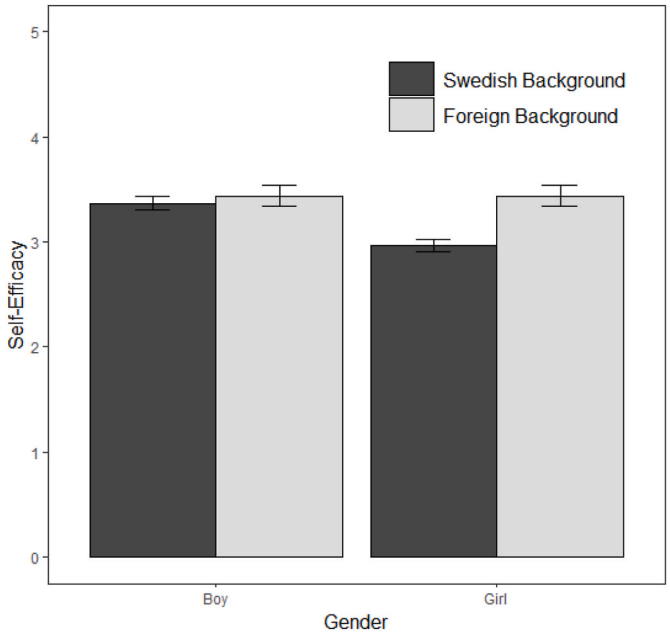

Fig. 2. Gender × Ethnicity Interaction Plot in Model Predicting Self-Efficacy (T1).

performance (Lin, 2016).

As predicted, both self-rated and teacher-rated performance was related to the participants self-efficacy ratings in this study. However, when we tested both factors on changes in self-efficacy development over time, we discovered an unexpected moderating effect of gender. First, only boys' development of programming self-efficacy over time was influenced by their self-rated performance in class and only girls' programming self-efficacy over time was related to the teacher-rated performance. Although we did not predict this result it fits with previous research which has shown that men in STEM tend to state that their interest developed from previous mastery experience, while women in STEM pointed at the importance of encouragement and teacher influence for their STEM-interest development (Zeldin et al., 2008).

4.2.2. Implications for future interventions

Since the programming education affected both boys' and girls' programming self-efficacy similarly, the gender difference remained at the end of the school year. This implies that to reduce the gender difference in programming self-efficacy, we need additional interventions to teaching programming in school. For girls' programming self-efficacy to fully benefit from their mastery experience, we likely need to counteract negative stereotypes of girls' technical ability, by for example educating children and the public of gender similarity in abilities. Encouraging specifically girls in STEM-classes also appears important, especially since their self-efficacy development relate to teacher's perceptions of their ability. Also, as we will discuss next, increasing self-efficacy is not always sufficient to increase interest.

4.3. The programming education reduced interest

According to EEVT (Eccles, 1987; 1994) and SCCT (Lent et al., 1994), self-efficacy is an important predictor of career interest, and much research has shown that gender differences in self-efficacy mediates gender differences in STEM-interest (e.g., Eccles & Wang, 2016; Hackett, 1995; Lent & Brown, 2019; Master & Meltzoff, 2020; Tellhed et al., 2017, 2018; Rottinghaus et al., 2003; Watt, 2010). As expected, we also found a gender difference in programming interest in this study. This result is coherent with a previous study with 14-year-old children in Hong Kong (Kong et al., 2018) and a study with 6-year-old children in the US (Master et al., 2017). Also, as expected, gender differences in self-efficacy completely mediated the gender difference in interest, through the three emerging sources of self-efficacy in this study: mastery experience, social persuasion, and vicarious experience.

Since the programming education was related to increases in the children's self-efficacy, and self-efficacy and interest were related, we were surprised to find that the participants' programming interest was *reduced* (not increased), after they took part in the programming education. This unexpected result importantly points to the fact that although self-efficacy is a strong predictor of interest, it is not the only predictor, which has previously been stated (Wang & Degol, 2013). Since the education is new, investigations should explore the children's perception of it more closely. As previously mentioned, many teachers assigned to teach programming because of this reform feel insecure in education methods and report a lack of teaching material (Official reports of the Swedish Government, 2021:70; The National Union of Teachers in TheNational Union of Teachers in Sweden, 2020). It has previously been suggested that making education fun is important to capture children's interest and deepen their learning (Hagenauer & Hascher, 2014). Perhaps some inspiration for teaching developments can be taken from the study of Master et al. (2017), where 6-year-old girls programmed computers in the laboratory and were more interested in programming afterwards. However, considerations should be taken to the children's ages when constructing teaching materials that students will enjoy. Programming toy robots could perhaps be considered "childish" by most 12-year-olds. Note that the children in our study rated programming as averagely interesting (on the scale). Since we did not measure interest in subjects such as math or history, we do not know if the students like regular school subjects better than programming. However, it is concerning that the children's interest decreased after the lessons, and it calls for further exploration of how this education can be made more stimulating for this age group.

4.4. Intersectional effects and the gender-equality paradox

Lastly, we performed an intersectional analysis to explore if participant gender interacted with participant ethnicity in programming self-efficacy and interest. The results showed a novel finding: Only girls with a Swedish background had lower self-efficacy for, programming as compared to the boys. Interestingly, girls with a foreign background did not rate self-efficacy lower than the boys. Further, the children with a Swedish background rated lower programming interest than the children with a foreign background.

Although we did not expect this result, it could perhaps relate to the gender equality paradox, which we mentioned in the introduction. This

paradox entails that research has found larger self-rated psychological gender differences in countries which rate higher on gender-equality indexes (e.g., Stoet & Geary, 2018). We did not ask the children in our study what country they and their parents were born in, but since Sweden is rated as the most gender-equal country in the EU (European Institute for Gender Equality, 2021), chances are that the children with a foreign background had roots in a country which is rated as lower in gender-equality. A study recently showed that the gender-equality paradox may be explained by more gender-equal countries having stronger gender stereotypes than other countries (Breda et al., 2020). This perhaps implies that girls with a Swedish background experience more stereotype threat in relation to technology, as compared to girls with a foreign background in Sweden. Future research should investigate this possibility making sure to have more participants with different ethnic backgrounds than in the present study. The result is also interesting in relation to the previous study of children's programming self-efficacy in Hong Kong (Kong et al., 2018), which did not find a gender difference. This calls for more extensive future research of cross-country differences in programming self-efficacy, to see if the gender-equality paradox fits also this self-rated psychological difference.

4.5. Limitations

We want to bring up some additional limitations to those already addressed in this discussion. One important limitation is that the study design lacked a control group. This limitation was necessary, since all children in Sweden receive programming education, but we encourage replications in countries which currently lack elementary school programming education, to randomize children to intervention versus controls. Replications both in Sweden and in other countries is also needed to overcome limitations with external validity and to test hypotheses of cross-country variation in gender difference size, drawn from theory about the gender-equality paradox.

Control was also limited in this study by the variation in programming education implementation, since standards for how to teach programming is still under development in Sweden. More studies are needed that test for effects of more controlled, homogenous, programming interventions, such as those demonstrated at science centers aimed at children.

Since the children's programming interest unexpectedly decreased in this study, we especially encourage future studies to probe the children's more nuanced attitudes towards the programming education, perhaps using qualitative methods, to come up with ideas for how education may increase interest in this age group.

5. Conclusions

Very few studies have investigated programming self-efficacy in children and this study showed expected gender differences in Sweden, a country famous for its gender equality but where women are still stereotyped as low in technical skills, and where very few women have entered the IT sector. This study also expanded previous research by showing that the gender difference in programming self-efficacy was completely mediated by gender differences in the children's previous mastery experience, social persuasion, and vicarious experience with computers. Interestingly, only Swedish girls, and not girls with a foreign background, rated their self-efficacy lower as compared to the boys, which needs further exploration to see if it relates to the gender-equality paradox.

The study also contributed to research by demonstrating that Sweden's school reform, where computer programming is now on the schedule, increased the children's programming self-efficacy. That learning programming in school increases self-efficacy is important knowledge, but our results suggests that to reduce the gender difference in programming self-efficacy, we need to also counteract negative stereotypes of women's technical skills and encourage girls programming efforts.

Lastly, self-efficacy was strongly related to interest, but surprisingly programming education reduced the children's interest in programming despite increasing their self-efficacy. More research is therefore needed to find ways to increase also interest in programming. To attract more women to IT, we thus need to find ways to not only to make girls think "Sure I can code", but also "and I want to!".

Funding

This work was supported by the Swedish Research Council for Health, Working Life and Welfare [grant numbers 2019-00334].

Credit author statement

Una Tellhed: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Investigation, Writing – original draft;**Fredrik Björklund**: Conceptualization, Funding acquisition, Methodology Writing – original draft. **Kalle Kallio Strand**: Investigation, Data curation, Formal analysis, Writing – original draft.

Declaration of competing interest

None.

Acknowledgment

We would like to thank August Nilsson, Annie Sollvén, Oskar Bång and Tina Rastegar for helping us collect the data. We also thank Stefan Persson for designing the survey and Professor Martin Bäckström for statistical advice.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chb.2022.107370.

References

Aronson, J., & Inzlicht, M. (2004). The ups and downs of attributional ambiguity: Stereotype vulnerability and the academic self-knowledge of African American college students. *Psychological Science, 15*(12), 829–836. https://doi.org/10.1111/j.0956-7976.2004.00763.x

Askar, P., & Davenport, D. (2009). An investigation of factors related to self-efficacy for Java programming among engineering students. *The Turkish Online Journal of Educational Technology, 8*(1), 26–32.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1986). *Social foundations of thought & action: A social cognitive theory*. Upper Saddle River, NJ: Pearson Education.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bench, S. W., Lench, H. C., Miner, K., Flores, S. A., & Liew, J. (2015). Gender gaps in overestimation of math performance. *Sex Roles, 72*(11–12), 536–546. https://doi.org/10.1007/s11199-015-0486-9

Betz, N. E., & Hackett, G. (1981). The relationship of career-related self-efficacy expectations to perceived career options in college women and men. *Journal of Counseling Psychology, 28*(5), 399–410. https://doi.org/10.1037/0022-0167.28.5.399

Breda, T., Jouini, E., Napp, C., & Thebault, G. (2020). Gender stereotypes can explain the gender-equality paradox. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 117*(49), 31063–31069. https://doi.org/10.1073/pnas.2008704117

Brown, S. D., Tramayne, S., Hoxha, D., Telander, K., Fan, X., & Lent, R. W. (2008). Social cognitive predictors of college students' academic performance and persistence: A meta-analytic path analysis. *Journal of Vocational Behavior, 72*(3), 298–308. https://doi-org.ludwig.lub.lu.se/10.1016/j.jvb.2007.09.003.

Cadaret, M. C., Hartung, P. J., Subich, L. M., & Weigold, I. K. (2017). Stereotype threat as a barrier to women entering engineering careers. *Journal of Vocational Behavior, 99*, 40–51. https://doi-org.ludwig.lub.lu.se/10.1016/j.jvb.2016.12.002.

Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in

STEM? *Social Psychological and Personality Science, 2*(6), 656–664. https://doi-org.ludwig.lub.lu.se/10.1177/1948550611405218.
Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin, 143*(1), 1–35. https://doi.org/10.1037/bul0000052
Cooper, J. (2006). The digital divide: The special case of gender. *Journal of Computer Assisted Learning, 22*(5), 320–334. https://doi-org.ludwig.lub.lu.se/10.1111/j.1365-2729.2006.00185.x.
Dennehy, T. C., & Dasgupta, N. (2017). Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences of the United States of America, 114*(23), 5964–5969. https://doi.org/10.1073/pnas.1613117114
Department of Higher Education. (2021). Tabell 2a. Sökande till gymnasieskolan till läsåret 2020/2021, efter kön [Table 2a. Applicants to high school 2020/2021, sorted by sex]. https://siris.skolverket.se/siris/sitevision_doc.getFile?p_id=550188. (Accessed 19 May 2022).
Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11*, 135–172. https://doi.org/10.1111/j.1471-6402.1987.tb00781.x
Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly, 18*(4), 585–609. https://doi.org/10.1111/j.1471-6402.1994.tb01049.x
Eccles, J. S., & Wang, M.-T. (2016). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development, 40*(2), 100–106. https://doi.org/10.1177/0165025415616201
Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*(1), 103–127. https://doi.org/10.1037/a0018053
Ertl, B., Luttenberger, S., & Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00703
European Institute for Gender Equality. (2017). *Gender segregation in education, training and the labour market. Review of the implementation of the Beijing Platform for Action in the EU Member States*. https://data.consilium.europa.eu/doc/document/ST-12709-2017-ADD-2/en/pdf. (Accessed 19 May 2022).
European Institute for Gender Equality. (2021). *Index score for European union for the 2021 edition*. https://eige.europa.eu/gender-equality-index/2021/country. (Accessed 19 May 2022).
Fessakis, G., Gouli, E., & Mavroudi, E. (2013). Problem solving by 5–6 years old kindergarten children in a computer programming environment: A case study. *Computers & Education, 63*, 87–97. https://doi-org.ludwig.lub.lu.se/10.1016/j.compedu.2012.11.016.
Fox, J., & Weisberg, S. (2019). *An {R} companion to applied regression* (3rd ed.). Thousand Oakes, CA: Sage https://socialsciences.mcmaster.ca/jfox/Books/Companion/.
Hackett, G. (1995). Self-efficacy in career choice and development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 232–258). New York: Cambridge University Press.
Hagenauer, G., & Hascher, T. (2014). Early adolescents' enjoyment experienced in learning situations at school and its relation to student achievement. *Journal of Education and Training Studies, 2*, 20–30. https://doi.org/10.11114/jets.v2i2.254
Huyer, S. (2015). Is the gender gap narrowing in science and engineering?. In F. Schlegel (Ed.), *UNESCO science report: Towards* (Vol. 2030, pp. 84–103). https://unesdoc.unesco.org/ark:/48223/pf0000235447. (Accessed 19 May 2022).
Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*, 581–592. https://doi.org/10.1037/0003-066X.60.6.581
Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., van Anders, & S. M.. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist, 74*(2), 171–193. https://doi.org/10.1037/amp0000307
Kalelioğlu, F., & Gülbahar, Y. (2014). The effects of teaching programming via scratch on problem solving skills: A discussion from learners' perspective. *Informatics in Education, 13*(1), 33–50. https://doi.org/10.15388/infedu.2014.03
Keren, G., & Fridin, M. (2014). Kindergarten social assistive robot (KindSAR) for children's geometric thinking and metacognitive development in preschool education: A pilot study. *Computers in Human Behavior, 35*, 400–412. https://doi-org.ludwig.lub.lu.se/10.1016/j.chb.2014.03.009.
Knight, C. R., & Brinton, M. C. (2017). One egalitarianism or several? Two decades of gender-role attitude change in europe. *American Journal of Sociology, 122*(5), 1485–1532. https://doi.org/10.1086/689814
Kong, S.-C., Chiu, M. M., & Lai, M. (2018). A study of primary school students' interest, collaboration attitude, and programming empowerment in computational thinking education. *Computers & Education, 127*, 178–189. https://doi.org/10.1016/j.compedu.2018.08.026
Korkmaz, O., & Altun, H. (2014). A validity and reliability study of the attitude scale of computer programming learning (ASCOPL). *Mevlana International Journal of Education, 4*(1), 30–43.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13
Lent, R, & Brown, S (2006). On conceptualizing and assessing social cognitive constructs in career research: A measurement guide. *Journal of Career Assessment, 14*(1), 12–35. https://doi.org/10.1177/1069072705281364
Lent, R. W., & Brown, S. D. (2019). Social cognitive career theory at 25: Empirical status of the interest, choice, and performance models. *Journal of Vocational Behavior, 115*, 1–14. https://doi.org/10.1016/j.jvb.2019.06.004
Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45*(1), 79–122. https://doi.org/10.1006/jvbe.1994.1027
Lin, G.-Y. (2016). Self-efficacy beliefs and their sources in undergraduate computing disciplines: An examination of gender and persistence. *Journal of Educational Computing Research, 53*(4), 540–561. https://doi.org/10.1177/0735633115608440
Mair, P., & Wilcox, R. R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods, 52*, 464–488. https://doi.org/10.3758/s13428-019-01246-w
Maltese, A. V., & Tai, R. H. (2010). Eyeballs in the fridge: Sources of early interest in science. *International Journal of Science Education, 32*, 669–685. https://doi.org/10.1080/09500690902792385
Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. N. (2017). Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology, 160*, 92–106. https://doi.org/10.1016/j.jecp.2017.03.013
Master, A., & Meltzoff, A. N. (2020). Cultural stereotypes and sense of belonging contribute to gender gaps in STEM. *International Journal of Gender, Science & Technology, 12*(1), 152–198. http://genderandset.open.ac.uk/index.php/genderandset/article/view/674/1124.
Meissner, F., Grigutsch, L. A., Koranyi, N., Müller, F., & Rothermund, K. (2019). Predicting behavior with implicit measures: Disillusioning findings, reasonable explanations, and sophisticated solutions. *Frontiers in Psychology, 10*, 1–16. https://doi.org/10.3389/fpsyg.2019.02483
Mellström, U. (2004). Machines and masculine subjectivity: Technology as an integral part of men's life experiences. *Men and Masculinities, 6*(4), 368–382. https://doi.org/10.1177/1097184X03260960
Mellström, U. (2009). The intersection of gender, race and cultural boundaries, or why is computer science in Malaysia dominated by women? *Social Studies of Science, 39*(6), 885–907. https://doi.org/10.1177/0306312709334636
Migrationsinfo. (2020). *Hur många i sverige är födda i ett annat land? [How many in Sweden are born abroad?]*. https://www.migrationsinfo.se/fragor-och-svar/hur-manga-utrikes-fodda-sverige/. (Accessed 19 May 2022).
Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology, 107*, 631–644. https://doi.org/10.1037/edu0000005
Ministry of Education. (2017). [National digitalisation strategy for the school system]. nationell-digitaliseringsstrategi-for-skolvasendet.pdf (regeringen.se). *Nationell digitaliseringsstrategi för skolväsendet*. https://www.regeringen.se/4a9d9a/contentassets/00b3d9118b0144f6bb95302f3e08d11c/nationell-digitaliseringsstrategi-for-skolvasendet.pdf.
National Science Foundation. (2021). *Women, minorities, and persons with disabilities in science and engineering: 2021*. NSF - National Science Foundation. https://ncses.nsf.gov/pubs/nsf21321/report.
Nordic Council of Ministers. (2021). *Genusperspektiv på framtidens högteknologiska arbetsliv: En nordisk forskningsöversikt om utbildningsval inom STEM (science, tehnology, engineering and Mathematics). [A gender perspective on the high-tech work life of the future: A nordic literature review of educational choice of STEM*. https://www.norden.org/sv/publication/genusperspektiv-pa-framtidens-hogteknologiska-arbetsliv. (Accessed 19 May 2022).
OECD. (2019). *OECD employment outlook 2019: The future of work*. Paris: OECD Publishing. https://doi.org/10.1787/9ee00155-en
Official reports of the Swedish Government (SOU 2021: 70). (2021). *Läromedelsutredningen: Böckernas betydelse och elevernas tillgång till kunskap. [Investigation of teaching materials: The significance of books and school pupils' access to knowledge]* (p. 70). https://www.regeringen.se/4a37bb/contentassets/e13e110fdc30401f9bf4a6ee8fa160a7/laromedelsutredningen–bockernas-betydelse-och-elevernas-tillgang-till-kunskap-sou-202170. (Accessed 19 May 2022).
Oldenziel, R. (1999). *Making technology masculine: Men, women, and modern machines in America, 1870–1945*. Amsterdam: Amsterdam University Press.
Olsson, M., & Martiny, S. E. (2018). Does exposure to counterstereotypical role models influence girls' and women's gender stereotypes and career choices? A review of social psychological research. *Frontiers in Psychology, 9*, 1–15. https://doi-org.ludwig.lub.lu.se/10.3389/fpsyg.2018.02264.
Phan, H. P., & Ngu, B. H. (2016). Sources of self-efficacy in academic contexts: A longitudinal perspective. *School Psychology Quarterly, 31*(4), 548–564. https://doi.org/10.1037/spq0000151
R Core Team. (2021). *R: A language and environment for statistical computing [computer software]*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/. (Accessed 19 May 2022).
Ramalingam, V., & Wiedenbeck, S. (1998). Development and validation of scores on a computer programming self-efficacy scale and group analyses of novice programmer self-efficacy. *Journal of Educational Computing Research, 19*(4), 367–381. https://doi.org/10.2190/C670-Y3C8-LTJ1-CT3P
Rogozhkina, I., & Kushnirenko, A. (2011). PiktoMir: Teaching programming concepts to preschoolers with a new tutorial environment. *Procedia - Social and Behavioral Sciences, 28*, 601–605. https://doi-org.ludwig.lub.lu.se/10.1016/j.sbspro.2011.11.114.
Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. http://www.jstatsoft.org/v48/i02/.
Rottinghaus, P. J., Larson, L. M., & Borgen, F. H. (2003). The relation of self-efficacy and interests: A meta-analysis of 60 samples. *Journal of Vocational Behavior, 62*, 221–236. https://doi.org/10.1016/S0001-8791(02)00039-8
Shoffner, M. F., & Dockery, D. J. (2015). Promoting interest in and entry into science, technology, engineering, and mathematics careers. In *APA handbook of career*

intervention, volume 2: Applications (pp. 125–137). Washington, DC: American Psychological Association.
Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology, 67*, 415–437. https://doi.org/10.1146/annurev-psych-073115-103235
Starr, C. R. (2018). "I'm not a science nerd!": STEM stereotypes, identity, and motivation among undergraduate women. *Psychology of Women Quarterly, 42*(4), 489–503. https://doi.org/10.1177/0361684318793848
Statistics Sweden. (2002). Statistics on persons with foreign background: Guidelines and recommendations. https://www.scb.se/contentassets/60768c27d88c434a8036d1fdb595bf65/mis-2002-3.pdf. (Accessed 19 May 2022).
Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613–629. https://doi.org/10.1037/0003-066X.52.6.613
Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797
Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological Science, 29*(4), 581–593. https://doi.org/10.1177/0956797617741719
Su, R., & Rounds, J. (2015). All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields. *Frontiers in Psychology, 6*(189), 1–20. https://doi.org/10.3389/fpsyg.2015.00189
Swedish Agency for Economic and Regional Growth. (2020). *Utmaningar och åtgärder för ökad jämställdhet inom IT-sektorn: En analys av varför kvinnor är underrepresenterade inom IT-sektorn och hur det kan förändras. [Challenges and interventions for increased gender equality in the IT-sector: An analysis of why women are under-represented in the IT-sector and how it can change]*. https://tillvaxtverket.se/download/18.2ef0f07117c0c64c87b34d44/1633090962537/Utmaningar%20och%20%C3%A5tg%C3%A4rder%20f%C3%B6r%20%C3%B6kad%20j%C3%A4mst%C3%A4lldhet%20inom%20IT-sektorn.pdf. (Accessed 19 May 2022).
Swedish Higher Education Authority. (2020). Higher education institutions in Sweden. 2020 status report. https://english.uka.se/download/18.7b31ebea172ea978184158b1/1598456063682/20-0102%20UKA%CC%88%202020%20Status%20Report.pdf. (Accessed 19 May 2022).
Swedish Media Council. (2021). *Ungar & medier 2021: En statistisk undersökning av ungas mediavanor och attityder till mediaanvändning[Kids and media 2021: A statistical investigation of young people's media habits and attitudes to media use]*. https://www.statensmedierad.se/download/18.57e3b0a17a75fbf0654916e/1632377641632/Ungar%20och%20medier%202021_anpassad.pdf. (Accessed 19 May 2022).
Swedish National Agency for Education. (2018). Curriculum for the compulsory school, preschool class and school-age educare. Revised 2018 https://www.skolverket.se/getFile?file=3984. (Accessed 16 June 2022).
Swedish IT and Telecom Industries. (2021). *Tech- Sweden's new basic industry: A report from Swedish it and telecom industries*. https://www.almega.se/app/uploads/sites/2/2021/05/ittelekomforetagen-tech-sveriges-nya-basidustri-2021-eng.pdf. (Accessed 19 May 2022).
Tek, F. B., Benli, K. S., & Deveci, E. (2018). Implicit theories and self-efficacy in an introductory programming course. *IEEE Transactions on Education, 61*(3), 218–225. https://doi.org/10.1109/TE.2017.2789183
Tellhed, U., & Adolfsson, C. (2018). Competence and confusion: How stereotype threat can make you a bad judge of your competence. *European Journal of Social Psychology, 48*(2), 189–197. https://doi.org/10.1002/ejsp.2307
Tellhed, U., Bäckström, M., & Björklund, F. (2017). Will I fit in and do well? The importance of social belongingness and self-efficacy for explaining gender differences in interest in STEM- and HEED-majors. *Sex Roles, 77*, 86–96. https://doi.org/10.1007/s11199-016-0694-y
Tellhed, U., Bäckström, M., & Björklund, F. (2018). The role of ability beliefs and agentic vs.communal career goals in adolescents first educational choice. What explains the degree of gender-balance? *Journal of Vocational Behavior, 104*, 1–13. https://doi.org/10.1016/j.jvb.2017.09.008
Tellhed, U., Björklund, F., & Kallio Strand, K. (2021). *"Technology is for boys, caregiving for girls". Strong implicit "people"-"things" stereotypes relate to Swedish children's STEM-interest*. Manuscript submitted for publication. Department of Psychology, (Lund University).
The National Union of Teachers in Sweden. (2020). *Programmering-en skolreform utan program. En lägesbeskrivning efter två år med obligatorisk undervisning i programmering. [Programming-a school reform which lacks a program. A status report after two years of obligatory programming teaching]*. https://www.lr.se/download/18.48c9289d171c2357975334/1588156019458/Programmering_En_skolreform_utan_program_LRUND182_202004.pdf.
Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software, 59*(5), 1–38. http://www.jstatsoft.org/v59/i05/.
Tsai, C.-Y. (2019). Improving students' understanding of basic programming concepts through visual programming language: The role of self-efficacy. *Computers in Human Behavior, 95*, 224–232. https://doi-org.ludwig.lub.lu.se/10.1016/j.chb.2018.11.038.
UNESCO. (2017). Cracking the code: Girls' and women's education in science, technology, engineering and mathematics (STEM). *United Nations Educational, Scientific and Cultural Organization*, 1–85. https://unesdoc.unesco.org/ark:/48223/pf0000253479.
Unrau, N. J., Rueda, R., Son, E., Polanin, R. J., Lundeen, R. J., & Muraszewski, A. K. (2018). Can reading self-efficacy be modified? A meta-analysis of the impact of interventions on reading self-efficacy. *Review of Educational Research, 88*(2), 167–204. https://doi.org/10.3102/0034654317743199
US Bureau of Labor Statistics. (2021). *Women in the labor force: A databook*. BLS Reports. U.S. Bureau of Labor Statistics.
Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751–796. https://doi-org.ludwig.lub.lu.se/10.3102/0034654308321456.
Usher, E. L., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. *Contemporary, Educational Psychology, 34*(1), 89–101. https://doi.org/10.1016/j.cedpsych.2008.09.002
Vermeulen, L., Van Bauwel, S., & Van Looy, J. (2017). Tracing female gamer identity. An empirical study into gender and stereotype threat perceptions. *Computers in Human Behavior, 71*, 90–98. https://doi-org.ludwig.lub.lu.se/10.1016/j.chb.2017.01.054.
Vinnervik, P. (2022a). Implementing programming in school mathematics and technology: Teachers' intrinsic and extrinsic challenges. *International Journal of Technology and Design Education, 32*, 213–242. https://doi.org/10.1007/s10798-020-09602-0
Vinnervik, P. (2022b). An in-depth analysis of programming in the Swedish school curriculum—rationale, knowledge content and teacher guidance. *Journal of Computers in Education*, 1–35. https://doi.org/10.1007/s40692-022-00230-2. Advance online publication.
Wang, M.-T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review, 33*, 304–340. https://doi-org.ludwig.lub.lu.se/10.1016/j.dr.2013.08.001.
Watt, H. G. (2010). Gender and occupational choice. In J. C. Chrisler, & D. R. McCreary (Eds.), *Handbook of gender research in psychology*. New York: Springer.
Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi-org.ludwig.lub.lu.se/10.1006/ceps.1999.1015.
Zeldin, A. L., Britner, S. L., & Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *Journal of Research in Science Teaching, 45*, 1036–1058. https://doi.org/10.1002/tea.20195
Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist, 70*(1), 10–20. https://doi.org/10.1037/a0038208.supp